

Doctorant·e en Sociologie et Informatique

Intitulé du poste

Doctorant·e en Sociologie et Informatique

Informations générales

Intitulé de l'offre : Doctorant·e en Sociologie et Informatique (H/F)

Référence : TODO

Lieu de travail : Lille

Responsable scientifique : Fabien ELOIRE, Maxime MORGE

Date de publication : 14 avril 2024

Type de contrat : CDD Doctorant/Contrat doctoral

Durée du contrat : 36 mois

Date de début de la thèse : 1 octobre 2024

Quotité de travail : Temps complet

Rémunération : 2 135,00 € brut mensuel

Sections CN : Sociologie et sciences du droit, Sciences de l'information : signaux, images, langues, automatique, robotique, interactions, systèmes intégrés matériel-logiciel

Description du sujet de thèse

Cette thèse s'inscrit dans le projet RORES-CL qui consiste à élaborer un modèle multi-agents explicatif des phénomènes sociaux qui se déploient sur les plateformes numériques publiques de questions-réponses telles que Stack Overflow. Adoptant une approche interdisciplinaire mêlant sociologie et informatique, ce projet cherche à appréhender les dynamiques sociales qui résultent de l'observation empirique des comportements individuels dans les communautés en ligne. Les travaux du ou de la doctorant·e qui sera recruté·e sur le projet se focaliseront sur les communautés de *data analysts* qui partagent un intérêt pour les outils informatiques d'analyse statistique sur la plateforme Stack Overflow (e.g. R ou Pandas). Le programme privilégie l'exploration et la démarche inductive en vue de poser les bases de la théorisation et de la modélisation centrée « individus ».

Dans la mesure où la plateforme Stack Overflow est une organisation qui crée de l'action collective, elle peut être modélisée à travers le concept classique de rôle social (Linton, 1936 ; Merton, 1949). Comme les interactions génèrent à la fois les

rôles sociaux et le statut social à travers la reconnaissance de l'expertise individuelle et le développement du score de réputation, l'analyse des interdépendances relationnelles à l'aide d'un modèle multi-agents (Gilbert, 1994) est un enjeu fort de cette recherche doctorale pour dépasser le caractère désincarné des données numériques. Elle requiert une solide démarche empirique de construction des hypothèses à partir de l'observation quantitative des comportements sur la plateforme, préalable à tout parti pris théorique ou méthodologique (analyse de réseaux, simulations, système multi-agents). Dans le cadre d'une démarche sociologique ouverte, l'enrichissement des données disponibles dans le cadre du projet est donc une possibilité laissée à la libre appréciation du/de la doctorant-e (entretiens, observations, accès à d'autres archives).

Des travaux préliminaires et exploratoires ont abouti à un entrepôt de données massives qui donne à voir à la fois les comportements des individus à l'échelle micro et les dynamiques collectives à l'échelle macro, tandis que l'approche en termes de réseaux permet de repérer, à l'échelle méso, différents phénomènes complexes au sein de chaque communauté (Delarre et al, 2023). Le doctorant aura pour mission :

1. d'enrichir le jeu de données aux autres actions de régulation associées aux rôles de modération comme les votes, les commentaires, à d'autres communautés/langages et d'un point de vue diachronique ;
2. d'étendre les analyses descriptives et de réseau à ces jeux de données à partir des concepts et hypothèses issues des études bibliographiques menées en parallèle ;
3. de calibrer le modèle à partir de l'observation des résultats empiriques des analyses.

Références bibliographiques

- Ralph Linton. 1936. « The Study Of Man: An Introduction » D. Appleton-Century Company, New York.
- Robert K. Merton. 1949. « Social Theory and Social Structure ». Free Press.
- Nigel Gilbert, éd. « Simulating societies : the computer simulation of social phenomena ». UCL Press, 1994.
- Sébastien Delarre, Fabien Eloire, Antoine Nongaillard, Maxime Morge. 2023. « Modèle explicatif de la sécession des experts dans les communautés de pratiques ». 31èmes journées francophones sur les systèmes multi-agents (JFSMA), Strasbourg, France. pp.65-74.

Contexte de travail

La thèse s'inscrit dans le projet RORES-CL qui bénéficie du soutien financier du CNRS à travers les programmes interdisciplinaires de la MITI. Elle s'appuie sur la

complémentarité des expertises du CLERSÉ (Centre Lillois d'Études et de Recherches Sociologiques et Économiques) et du laboratoire CRISAL (Centre de Recherche en Informatique, Signal et Automatique de Lille) à l'Université de Lille.

Contraintes et risques

Cette thèse de doctorat sera co-dirigée par Fabien Eloire, professeur de sociologie et Maxime Morge, maître de conférences HDR d'informatique. Co-encadré·e par Sébastien Delarre (maître de conférences en sociologie), le/la doctorant·e sera inscrit·e à l'École doctorale SESAM de l'Université de Lille. Les réunions de thèse seront bimensuelles. Sa station de travail se situera dans le bâtiment ESPRIT au cœur du campus scientifique de Villeneuve-d'Ascq (Hauts-de-France) et disposera d'un accès à la plateforme informatique Data/UdL.

Informations complémentaires

Le candidat devra être titulaire d'un master en Sociologie ou d'un diplôme d'ingénieur en Informatique. Le poste nécessite une solide formation en théorie sociologique et en analyse de données, de bonnes aptitudes de communication orale et écrite — en français (C1) et en anglais (B2) — pour présenter aux congrès et rédiger des articles dans des revues scientifiques. Nous recherchons un·e jeune chercheur·se qui saura s'impliquer dans son projet, curieux·se, ayant une certaine autonomie et une forte motivation pour développer une triple compétence en sociologie des réseaux et en simulation multi-agents et en science des données. De plus, le candidat devra être apte à travailler en équipe dans un projet interdisciplinaire.

Les candidatures devront inclure :

- une lettre de motivation (une page minimum). Les lettres de motivation génériques seront ignorées ;
- un curriculum vitae (CV) ;
- des lettres de recommandation de deux personnes de référence ;
- le relevé de notes du master ou d'école d'ingénieur.

PhD thesis in Sociology and Computer Science

Job title

PhD thesis in Sociology and Computer Science

General information

Job title: PhD thesis in Sociology and Computer Science (M/F)

Reference : TODO

Place of work : Lille

Scientific supervisor: Fabien ELOIRE, Maxime MORGE

Publication date: 14th of April

Type of contract: Fixed-term doctoral contract

Contract duration: 36 months

Start date of thesis: 1 October 2024

Working hours: full time

Salary: €2,135.00 gross per month

Sections CN: Sociology and legal sciences; Information sciences: signals, images, languages, automation, robotics, interactions, hardware-software integrated systems

Description of thesis subject

This thesis is part of the RORES-CL project, which involves developing an agent-based model to explain the social phenomena that occur on public question-and-answer digital platforms such as Stack Overflow. Adopting an interdisciplinary approach combining sociology and computer science, this project seeks to understand the social dynamics that result from individual behaviour in online communities. The work of the PhD candidate will focus on communities of data analysts who share an interest in statistical analysis tools (e.g. R or Pandas) on the Stack Overflow platform. The programme emphasises exploration and an inductive approach in order to lay the foundations for theorising and agent-based modelling.

Since Stack Overflow is an organisation that creates collective action, it can be modelled using the classic concept of social role (Linton, 1936; Merton, 1949). As interactions generate both social roles and social status through the recognition of individual expertise and the development of reputation scores, the analysis of relational interdependencies using an agent-based model (Gilbert, 1994) is a major challenge for this doctoral research in order to overcome the disembodied nature of digital data. It requires a solid empirical approach for the construction of

hypotheses based on the quantitative observation of behaviour on the platform, prior to any theoretical or methodological bias (network analysis, agent-based modelling). As part of an open sociological approach, the doctoral student is free to enrich the data available as part of the project (interviews, observations, access to other archives).

Preliminary work has resulted in a massive data warehouse that focus on both the individual behaviours at the microscopic scale and the collective dynamics at the macroscopic scale. The network approach makes it possible to identify, at a mesoscopic scale, various complex phenomena within each community (Delarre et al, 2023). The PhD candidate will be responsible for:

1. enrich the dataset with other regulatory actions associated with moderation roles, such as voting and commenting, in other communities/languages and from a diachronic point of view;
2. extend the multivariate descriptive and social network analysis to these datasets on the basis of the concepts and hypotheses derived from the literature studies carried out in parallel;
3. calibrate the model based on observation of the empirical results of the analyses.

References

- Ralph Linton. 1936. « The Study Of Man: An Introduction » D. Appleton-Century Company, New York.
- Robert K. Merton. 1949. « Social Theory and Social Structure ». Free Press.
- Nigel Gilbert, éd. « Simulating societies: the computer simulation of social phenomena ». UCL Press, 1994.
- Sébastien Delarre, Fabien Eloire, Antoine Nongaillard, Maxime Morge. 2023. « Modèle explicatif de la sécession des experts dans les communautés de pratiques ». 31èmes journées francophones sur les systèmes multi-agents (JFSMA), Strasbourg, France. pp.65-74.

Working context

The thesis is part of the RORES-CL project, which receives financial support from the CNRS through the MITI interdisciplinary programmes. It is supported by the complementary expertise of CLERSÉ (Centre Lillois d'Études et de Recherches Sociologiques et Économiques) and the CRISAL laboratory (Centre de Recherche en Informatique, Signal et Automatique de Lille) at the University of Lille.

Constraints and risks

This PhD thesis will be co-directed by Fabien Eloire, professor of sociology, and Maxime Morge, associated professor in computer science. Co-supervised by Sébastien Delarre (associated professor in sociology), the PhD candidate will be enrolled in the SESAM doctoral school at the University of Lille. Meetings will be held twice a month. The workstation will be located in the ESPRIT building at the heart of the Villeneuve-d'Ascq science campus (Hauts-de-France) and the PdD candidate will have access to the Data/UdL IT platform.

Further information

The candidate should have a Master's degree in Sociology or an engineering degree in Computer Science. The job requires a solid background in sociological theory and data analysis, good oral and written communication skills — in French (C1) and English (B2) — for presenting at conferences and writing articles for scientific journals. We are looking for a young researcher who is committed to the project, curious, with a degree of autonomy and a strong motivation to develop to triple skills in the sociology of network, agent-based simulation and data science. In addition, the candidate should be able to work as part of a team on an interdisciplinary project.

Applications must include:

- a covering letter (minimum one page). Generic cover letters will be ignored;
- a curriculum vitae (CV)
- letters of recommendation from two referees;
- transcript of marks from master's degree or engineering school.